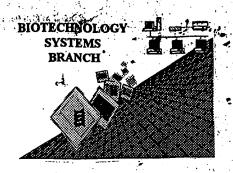# RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: **09/673,274**

Source: **Pu/09**

Date Processed by STIC: **8/1/2001**

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.**

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE **CHECKER VERSION 3.0 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

Checker Version 3.0 can be down loaded from the USPTO website at the following address:
http://www.uspto.gov/web/offices/pac/checker

ERROR DETECTED   **SUGGESTED CORRECTION**   SERIAL NUMBER: _09/673,274_

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1_____Wrapped Nucleics
       Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2_____Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3_____Misaligned Amino
       Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4_____Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5_____Variable Length.

Sequence(s)_____contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6_____PatentIn 2.0
       "bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7_____Skipped Sequences
       (OLD RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8_____Skipped Sequences
       (NEW RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9_____Use of n's or Xaa's
       (NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10____Invalid <213>
       Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11_____Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses.
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12_____PatentIn 2.0
       "bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13_____Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001

PCT09

delete xtra <110's - only

one <110 should be shown

```
          RAW SEQUENCE LISTING            DATE: 08/01/2001
          PATENT APPLICATION:  US/09/673,274    TIME: 19:08:03

              Input Set : A:\98015seq.txt
              Output Set: N:\CRF3\08012001\I673274.raw
```

Does Not Comply
Corrected Diskette Needed
JR 1-8

```
     3 <110> APPLICANT: LAMBERTY, MIREILLE
W--> 4 <110> APPLICANT: BULET, PHILIPPE
W--> 5 <110> APPLICANT: BROOKHART, GARY
W--> 6 <110> APPLICANT: HOFFMAN, JULES
     8 <120> TITLE OF INVENTION: GENE CODING FOR HELIOMICINE, AND USE THEREOF
    10 <130> FILE REFERENCE: A33595 PCT USA
    12 <140> CURRENT APPLICATION NUMBER: US 09/673,274
    13 <141> CURRENT FILING DATE: 1999-07-12
    15 <150> PRIOR APPLICATION NUMBER: PCT/FR99/00843
    16 <151> PRIOR FILING DATE: 1999-04-12
    18 <150> PRIOR APPLICATION NUMBER: FR 98 04933
    19 <151> PRIOR FILING DATE: 1998-04-15
    21 <160> NUMBER OF SEQ ID NOS: 38
    23 <170> SOFTWARE: PatentIn Ver. 2.1
```

ERRORED SEQUENCES

```
    25 <210> SEQ ID NO: 1
    26 <211> LENGTH: 147
E-->27 <212> TYPE: ADN  DNA        global error
    28 <213> ORGANISM: synthetic construct   see item 10 on Error Summary Sheet
    30 <220> FEATURE:
    31 <221> NAME/KEY: CDS
    32 <222> LOCATION: (1)..(147)
    34 <400> SEQUENCE: 1
    35 agc ttg gat aaa aga gac aag ttg att ggc agc tgt gtt tgg ggc gcc    48
    36 Ser Leu Asp Lys Arg Asp Lys Leu Ile Gly Ser Cys Val Trp Gly Ala
    37  1               5                  10                 15
    39 gtc aac tac act agt gac tgc aac ggc gag tgc aag cgc cgc ggt tac    96
    40 Val Asn Tyr Thr Ser Asp Cys Asn Gly Glu Cys Lys Arg Arg Gly Tyr
    41             20                 25                 30
    43 aag ggt ggc cat tgt gga tcc ttc gct aac gtt aac tgt tgg tgt gaa   144
    44 Lys Gly Gly His Cys Gly Ser Phe Ala Asn Val Asn Cys Trp Cys Glu
    45          35                 40                 45
    47 acc                                                              147
    48 Thr
    51 <210> SEQ ID NO: 2
    52 <211> LENGTH: 169
E-->53 <212> TYPE: ADN
    54 <213> ORGANISM: synthetic construct
    56 <220> FEATURE:
    57 <221> NAME/KEY: CDS
    58 <222> LOCATION: (1)..(132)
    60 <400> SEQUENCE: 2
    61 gat aag ctt atc ggt tcc tgc gtg tgg ggt gct gtg aac tac act tcc    48
    62 Asp Lys Leu Ile Gly Ser Cys Val Trp Gly Ala Val Asn Tyr Thr Ser
```

RAW SEQUENCE LISTING   DATE: 08/01/2001
PATENT APPLICATION: US/09/673,274   TIME: 19:08:03

Input Set : A:\98015seq.txt
Output Set: N:\CRF3\08012001\I673274.raw

```
63   1           5              10             15
65 gat tgc aac ggt gag tgc aag agg agg ggt tac aag ggt ggt cac tgc    96
66 Asp Cys Asn Gly Glu Cys Lys Arg Arg Gly Tyr Lys Gly Gly His Cys
67           20              25              30
69 ggt tcc ttc gct aac gtg aac tgc tgg tgc gag act tgagagctcg        142
70 Gly Ser Phe Ala Asn Val Asn Cys Trp Cys Glu Thr
71           35              40
73 gcgaggcgaa cgtgtcgacg gatccgg                                     169
76 <210> SEQ ID NO: 3
77 <211> LENGTH: 261
```
E--> 78 <212> TYPE: (ADN)
```
79 <213> ORGANISM: (synthetic construct)
81 <220> FEATURE:
82 <221> NAME/KEY: CDS
83 <222> LOCATION: (3)..(224)
85 <400> SEQUENCE: 3
86 cc atg ggt ttc gtg ctt ttc tct cag ctt cca tct ttc ctt ctt gtg    47
87    Met Gly Phe Val Leu Phe Ser Gln Leu Pro Ser Phe Leu Leu Val
88    1           5              10             15
90 tct act ctt ctt ctt ttc ctt gtg atc tct cac tct tgc cgt gcc gat   95
91 Ser Thr Leu Leu Leu Phe Leu Val Ile Ser His Ser Cys Arg Ala Asp
92           20              25              30
94 aag ctt atc ggt tcc tgc gtg tgg ggt gct gtg aac tac act tcc gat  143
95 Lys Leu Ile Gly Ser Cys Val Trp Gly Ala Val Asn Tyr Thr Ser Asp
96           35              40              45
98 tgc aac ggt gag tgc aag agg agg ggt tac aag ggt ggt cac tgc ggt  191
99 Cys Asn Gly Glu Cys Lys Arg Arg Gly Tyr Lys Gly Gly His Cys Gly
100          50              55              60
102 tcc ttc gct aac gtg aac tgc tgg tgc gag act tgagagctcg gcgaggcgaa 244
103 Ser Phe Ala Asn Val Asn Cys Trp Cys Glu Thr
104          65              70
106 cgtgtcgacg gatccgg                                               261
109 <210> SEQ ID NO: 4
110 <211> LENGTH: 120
```
E--> 111 <212> TYPE: (ADN)
```
112 <213> ORGANISM: (synthetic construct)
114 <220> FEATURE:
115 <221> NAME/KEY: CDS
116 <222> LOCATION: (12)..(101)
118 <400> SEQUENCE: 4
119 gcgtcgacgc g atg ggt ttc gtg ctt ttc tct cag ctt cca tct ttc ctt  50
120             Met Gly Phe Val Leu Phe Ser Gln Leu Pro Ser Phe Leu
121             1           5              10
123 ctt gtg tct act ctt ctt ctt ttc ctt gtg atc tct cac tct tgc cgt   98
124 Leu Val Ser Thr Leu Leu Leu Phe Leu Val Ile Ser His Ser Cys Arg
125     15              20              25
127 gct ggagacgcga attcacaca                                          120
128 Ala
129 30
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/673,274

DATE: 08/01/2001
TIME: 19:08:03

Input Set : A:\98015seq.txt
Output Set: N:\CRF3\08012001\I673274.raw

```
280 <220> FEATURE:
281 <221> NAME/KEY: CDS
282 <222> LOCATION: (7)..(204)
284 <400> SEQUENCE: 18
285 tctaga atg gcc tgc acc aac aac gcc atg agg gcc ctc ttc ctc ctc      48
286        Met Ala Cys Thr Asn Asn Ala Met Arg Ala Leu Phe Leu Leu
287         1               5                  10
289 ctg ctc ttc tgc atc gtg cac ggc gat aag ctt atc ggt tcc tgc gtg     96
290 Leu Leu Phe Cys Ile Val His Gly Asp Lys Leu Ile Gly Ser Cys Val
291  15                20                  25                30
293 tgg ggt gct gtg aac tac act tcc gat tgc aac ggt gag tgc aag agg    144
294 Trp Gly Ala Val Asn Tyr Thr Ser Asp Cys Asn Gly Glu Cys Lys Arg
295                 35                  40                  45
297 agg ggt tac aag ggt ggt cac tgc ggt tcc ttc gct aac gtg aac tgc    192
298 Arg Gly Tyr Lys Gly Gly His Cys Gly Ser Phe Ala Asn Val Asn Cys
299            50                  55                  60
301 tgg tgc gag act tgactcgag
302 Trp Cys Glu Thr                                                    213
303            65
306 <210> SEQ ID NO: 19
307 <211> LENGTH: 838
308 <212> TYPE: ADN
309 <213> ORGANISM: synthetic construct
311 <220> FEATURE:
312 <221> NAME/KEY: promoter
313 <222> LOCATION: (7)..(532)
315 <220> FEATURE:
316 <221> NAME/KEY: misc_structure
317 <222> LOCATION: (533)..(568)
319 <220> FEATURE:
320 <221> NAME/KEY: terminator
321 <222> LOCATION: (569)..(832)
323 <400> SEQUENCE: 19
324 aagcttccag aaggtaatta tccaagatgt agcatcaaga atccaatgtt tacgggaaaa 60
326 actatggaag tattatgtga gctcagcaag aagcagatca atatgcggca catatgcaac 120
328 ctatgttcaa aaatgaagaa tgtacagata caagatccta tactgccaga atacgaagaa 180
330 gaatacgtag aaattgaaaa agaagaacca ggcgaagaaa agaatcttga agacgtaagc 240
332 actgacgaca acaatgaaaa gaagaagata ggtcggtga ttgtgaaaga gacatagagg 300
334 acacatgtaa ggtggaaaat gtaagggcgg aggtcggtga ttgtgaaaga gacatagagg 300
336 ccccactact tatccttta tattttccg tgtcattttt gcccttgagt tttcctatat 420
338 aaggaaccaa gttcggcatt tgtgaaaaca agaaaaaatt tggtgtaagc tattttcttt 480
340 gaagtactga ggatacaact tcagagaaat ttgtaagttt gtagatctcg attctagaag 540
342 gcctgaattc gagctcggta ccggatccaa ttcccgatcg ttcaaacatt tggcaataaa 600
344 gtttcttaag attgaatcct gttgccggtc ttgcgatgat tatcatataa tttctgttga 660
346 attacgttaa gcatgtaata attaacatgt aatgcatgac gttattttatg agatgggttt 720
348 ttatgattag agtcccgcaa ttatacattt aatacgcgat agaaaacaaa atatagcgcg 780
350 caaactagga taaattatcg cgcgcggtgt catctatgtt actagatcgg ggatcgat    838
353 <210> SEQ ID NO: 20
354 <211> LENGTH: 1036
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/673,274
DATE: 08/01/2001
TIME: 19:08:03

Input Set : A:\98015seq.txt
Output Set: N:\CRF3\08012001\I673274.raw

```
    132 <210> SEQ ID NO: 5
    133 <211> LENGTH: 75
E--> 134 <212> TYPE: ADN
    135 <213> ORGANISM: synthetic construct
    137 <400> SEQUENCE: 5
    138 gcgtcgacgc gatgggtttc gtgctttct ctcagcttcc atctttcctt cttgtgtcta 60
    140 ctcttcttct tttcc
    143 <210> SEQ ID NO: 6                                                 75
    144 <211> LENGTH: 72
E--> 145 <212> TYPE: ADN
    146 <213> ORGANISM: synthetic construct
    148 <400> SEQUENCE: 6
    149 tcgccggcac ggcaagagta agagatcaca aggaaaagaa gaagagtaga cacaagaagg 60
    151 aaagatggaa gc
    154 <210> SEQ ID NO: 7                                                 72
    155 <211> LENGTH: 80
E--> 156 <212> TYPE: ADN
    157 <213> ORGANISM: synthetic construct
    159 <400> SEQUENCE: 7
    160 gataagctta tcggttcctg cgtgtggggt gctgtgaact acacttccga ttgcaacggt 60
    162 gagtgcaaga ggaggggtta
    165 <210> SEQ ID NO: 8                                                 80
    166 <211> LENGTH: 109
E--> 167 <212> TYPE: ADN
    168 <213> ORGANISM: synthetic construct
    170 <400> SEQUENCE: 8
    171 ccggatccgt cgacacgttc gcctcgccga gctctcaagt ctcgcaccag cagttcacgt 60
    173 tagcgaagga accgcagtga ccaccttgt aacccctcct cttgcactc             109
    176 <210> SEQ ID NO: 9
    177 <211> LENGTH: 85
E--> 178 <212> TYPE: ADN
    179 <213> ORGANISM: synthetic construct
    181 <400> SEQUENCE: 9
    182 agggcccct agggtttaaa cggccagtca ggccgaattc gagctcggta cccgggggatc 60
    184 ctctagagtc gacctgcagg catgc
    187 <210> SEQ ID NO: 10                                                85
    188 <211> LENGTH: 66
E--> 189 <212> TYPE: ADN
    190 <213> ORGANISM: synthetic construct
    192 <400> SEQUENCE: 10
    193 ccctgaacca ggctcgaggg cgcgccttaa ttaaaagctt gcatgcctgc aggtcgactc 60
    195 tagagg
    198 <210> SEQ ID NO: 11                                                66
    199 <211> LENGTH: 93
E--> 200 <212> TYPE: ADN
    201 <213> ORGANISM: synthetic construct
    203 <400> SEQUENCE: 11
    204 ccggccagtc aggccacact taattaagtt taaacgcggc cccggcgcgc ctaggtgtgt 60
    206 gctcgagggc ccaacctcag tacctggttc agg
                                                                          93
```

RAW SEQUENCE LISTING   DATE: 08/01/2001
PATENT APPLICATION: US/09/673,274   TIME: 19:08:03

Input Set : A:\98015seq.txt
Output Set: N:\CRF3\08012001\I673274.raw

```
      209 <210> SEQ ID NO: 12
      210 <211> LENGTH: 93
E--> 211 <212> TYPE: ADN
      212 <213> ORGANISM: synthetic construct
      214 <400> SEQUENCE: 12
      215 ccggcctgaa ccaggtactg aggttgggcc ctcgagcaca cacctaggcg cgccgggggcc 60
      217 gcgtttaaac ttaattaagt gtggcctgac tgg                              93
      220 <210> SEQ ID NO: 13
      221 <211> LENGTH: 50
E--> 222 <212> TYPE: ADN
      223 <213> ORGANISM: synthetic construct
      225 <400> SEQUENCE: 13
      226 ggtctagaat ggcctgcacc aacaacgcca tgagggccct cttcctcctc             50
      229 <210> SEQ ID NO: 14
      230 <211> LENGTH: 50
E--> 231 <212> TYPE: ADN
      232 <213> ORGANISM: synthetic construct
      234 <400> SEQUENCE: 14
      235 ccgaattcgg cgccgtgcac gatgcagaag agcacgagga ggaagagggc             50
      238 <210> SEQ ID NO: 15
      239 <211> LENGTH: 81
E--> 240 <212> TYPE: ADN
      241 <213> ORGANISM: synthetic construct
      243 <220> FEATURE:
      244 <221> NAME/KEY: CDS
      245 <222> LOCATION: (7)..(72)
      247 <400> SEQUENCE: 15
      248 tctaga atg gcc tgc acc aac aac gcc atg agg gcc ctc ttc ctc ctc     48
      249        Met Ala Cys Thr Asn Asn Ala Met Arg Ala Leu Phe Leu Leu
      250         1               5                   10
      252 ctg ctc ttc tgc atc gtg cac ggc gccgaattc                          81
      253 Leu Leu Phe Cys Ile Val His Gly
      254  15                  20
      257 <210> SEQ ID NO: 16
      258 <211> LENGTH: 24
E--> 259 <212> TYPE: ADN
      260 <213> ORGANISM: synthetic construct
      262 <400> SEQUENCE: 16
      263 gataagctta tcggttcctg cgtg                                         24
      266 <210> SEQ ID NO: 17
      267 <211> LENGTH: 32
E--> 268 <212> TYPE: ADN
      269 <213> ORGANISM: synthetic construct
      271 <400> SEQUENCE: 17
      272 ggctcgagtc aagtctcgca ccagcagttc ac                                32
      275 <210> SEQ ID NO: 18
      276 <211> LENGTH: 213
E--> 277 <212> TYPE: ADN
      278 <213> ORGANISM: synthetic construct
```

RAW SEQUENCE LISTING   DATE: 08/01/2001
PATENT APPLICATION: US/09/673,274   TIME: 19:08:03

Input Set : A:\98015seq.txt
Output Set: N:\CRF3\08012001\I673274.raw

```
E--> 355 <212> TYPE: ADN
     356 <213> ORGANISM: synthetic construct
     358 <220> FEATURE:
     359 <221> NAME/KEY: promoter
     360 <222> LOCATION: (7)..(532)
     362 <220> FEATURE:
     363 <221> NAME/KEY: CDS
     364 <222> LOCATION: (539)..(736)
     366 <220> FEATURE:
     367 <221> NAME/KEY: terminator
     368 <222> LOCATION: (767)..(1030)
     370 <400> SEQUENCE: 20
     371 aagcttccag aaggtaatta tccaagatgt agcatcaaga atccaatgtt tacgggaaaa 60
     373 actatggaag tattatgtga gctcagcaag aagcagatca atatgcggca catatgcaac 120
     375 ctatgttcaa aaatgaagaa tgtacagata caagatccta tactgccaga atacgaagaa 180
     377 gaatacgtag aaattgaaaa agaagaacca ggcgaagaaa agaatcttga agacgtaagc 240
     379 actgacgaca acaatgaaaa gaagaagata aggtcggtga ttgtgaaaga gacatagagg 300
     381 acacatgtaa ggtggaaaat gtaagggcgg aaagtaacct tatcacaaag gaatcttatc 360
     383 ccccactact tatccttta tattttccg tgtcatttt gcccttgagt tttcctatat 420
     385 aaggaaccaa gttcggcatt tgtgaaaaca agaaaaaatt tggtgtaagc tattttcttt 480
     387 gaagtactga ggatacaact tcagagaaat ttgtaagttt gtagatctcg attctaga 538
     389 atg gcc tgc acc aac aac gcc atg agg gcc ctc ttc ctc ctc gtg ctc 586
     390 Met Ala Cys Thr Asn Asn Ala Met Arg Ala Leu Phe Leu Leu Val Leu
     391  1               5                   10                  15
     393 ttc tgc atc gtg cac ggc gat aag ctt atc ggt tcc tgc gtg tgg ggt 634
     394 Phe Cys Ile Val His Gly Asp Lys Leu Ile Gly Ser Cys Val Trp Gly
     395              20                  25                  30
     397 gct gtg aac tac act tcc gat tgc aac ggt gag tgc aag agg agg ggt 682
     398 Ala Val Asn Tyr Thr Ser Asp Cys Asn Gly Glu Cys Lys Arg Arg Gly
     399              35                  40                  45
     401 tac aag ggt ggt cac tgc ggt tcc ttc gct aac gtg aac tgc tgg tgc 730
     402 Tyr Lys Gly Gly His Cys Gly Ser Phe Ala Asn Val Asn Cys Trp Cys
     403          50                  55                  60
     405 gag act tgactcgagg ggggcccgg taccggatcc aattcccgat cgttcaaaca 786
     406 Glu Thr
     407  65
     409 tttggcaata aagtttctta agattgaatc ctgttgccgg tcttgcgatg attatcatat 846
     411 aatttctgtt gaattacgtt aagcatgtaa taattaacat gtaatgcatg acgttatta 906
     413 tgagatgggt ttttatgatt agagtcccgc aattatacat ttaatacgcg atagaaaaca 966
     415 aaatatagcg cgcaaactag gataaattat cgcgcgcggt gtcatctatg ttactagatc 1026
     417 ggggatcgat 1036
     420 <210> SEQ ID NO: 21
     421 <211> LENGTH: 52
E--> 422 <212> TYPE: ADN
     423 <213> ORGANISM: synthetic construct
     425 <400> SEQUENCE: 21
     426 agcttggata aaagagacaa gttgattggc agctgtgttt ggggcgccgt ca 52
     429 <210> SEQ ID NO: 22
     430 <211> LENGTH: 56
```

RAW SEQUENCE LISTING   DATE: 08/01/2001
PATENT APPLICATION: US/09/673,274   TIME: 19:08:03

Input Set : A:\98015seq.txt
Output Set: N:\CRF3\08012001\I673274.raw

```
E--> 431 <212> TYPE: ADN
     432 <213> ORGANISM: synthetic construct
     434 <400> SEQUENCE: 22
     435 agtgtagttg acggcgcccc aaacacagct gccaatcaac ttgtctcttt tatcca      56
     438 <210> SEQ ID NO: 23
     439 <211> LENGTH: 52
E--> 440 <212> TYPE: ADN
     441 <213> ORGANISM: synthetic construct
     443 <400> SEQUENCE: 23
     444 actacactag tgactgcaac ggcgagtgca agcgccgcgg ttacaagggt gg      52
     447 <210> SEQ ID NO: 24
     448 <211> LENGTH: 52
E--> 449 <212> TYPE: ADN
     450 <213> ORGANISM: synthetic construct
     452 <400> SEQUENCE: 24
     453 cacaatggcc acccttgtaa ccgcggcgct tgcactcgcc gttgcagtca ct      52
     456 <210> SEQ ID NO: 25
     457 <211> LENGTH: 56
E--> 458 <212> TYPE: ADN
     459 <213> ORGANISM: synthetic construct
     461 <400> SEQUENCE: 25
     462 ccattgtgga tccttcgcta acgttaactg ttggtgtgaa acctgatagg tcgaca      56
     465 <210> SEQ ID NO: 26
     466 <211> LENGTH: 52
E--> 467 <212> TYPE: ADN
     468 <213> ORGANISM: synthetic construct
     470 <400> SEQUENCE: 26
     471 gatctgtcga cctatcaggt ttcacaccaa cagttaacgt tagcgaagga tc      52
     474 <210> SEQ ID NO: 27
     475 <211> LENGTH: 42
E--> 476 <212> TYPE: ADN
     477 <213> ORGANISM: synthetic construct
     479 <400> SEQUENCE: 27
     480 gatccttcgc taacgttaac tgttggtgta gaacctgata gg      42
     483 <210> SEQ ID NO: 28
     484 <211> LENGTH: 42
E--> 485 <212> TYPE: ADN
     486 <213> ORGANISM: synthetic construct
     488 <400> SEQUENCE: 28
     489 tcgacctatc aggttctaca ccaacagtta acgttagcga ag      42
     492 <210> SEQ ID NO: 29
     493 <211> LENGTH: 32
E--> 494 <212> TYPE: ADN
     495 <213> ORGANISM: synthetic construct
     497 <400> SEQUENCE: 29
     498 ctagtgactg caacggcgag tgcttgttgc gc      32
     501 <210> SEQ ID NO: 30
     502 <211> LENGTH: 26
E--> 503 <212> TYPE: ADN
```

```
    504 <213> ORGANISM: synthetic construct
    506 <400> SEQUENCE: 30
    507 gcaacaagca ctcgccgttg cagtca                              26
    510 <210> SEQ ID NO: 31
    511 <211> LENGTH: 32
E--> 512 <212> TYPE: ADN
    513 <213> ORGANISM: synthetic construct
    515 <400> SEQUENCE: 31
    516 ctagtgactg cgctgctgag tgcaagcggc gc                       32
    519 <210> SEQ ID NO: 32
    520 <211> LENGTH: 26
E--> 521 <212> TYPE: ADN
    522 <213> ORGANISM: synthetic construct
    524 <400> SEQUENCE: 32
    525 gccgcttgca ctcagcagcg cagtca                              26
    528 <210> SEQ ID NO: 33
    529 <211> LENGTH: 40
E--> 530 <212> TYPE: ADN
    531 <213> ORGANISM: synthetic construct
    533 <400> SEQUENCE: 33
    534 agcttggata aaagagctgc tgctgctggt agctgtgttt               40
    537 <210> SEQ ID NO: 34
    538 <211> LENGTH: 18
E--> 539 <212> TYPE: ADN
    540 <213> ORGANISM: synthetic construct
    542 <400> SEQUENCE: 34
    543 ggggcgccgt caactaca                                       18
    546 <210> SEQ ID NO: 35
    547 <211> LENGTH: 22
E--> 548 <212> TYPE: ADN
    549 <213> ORGANISM: synthetic construct
    551 <400> SEQUENCE: 35
    552 ctagtgtagt tgacggcgcc cc                                  22
    555 <210> SEQ ID NO: 36
    556 <211> LENGTH: 36
E--> 557 <212> TYPE: ADN
    558 <213> ORGANISM: synthetic construct
    560 <400> SEQUENCE: 36
    561 aaacacagct accagcagca gcagctcttt tatcca                   36
    564 <210> SEQ ID NO: 37
    565 <211> LENGTH: 32
E--> 566 <212> TYPE: ADN
    567 <213> ORGANISM: synthetic construct
    569 <400> SEQUENCE: 37
    570 ctagtgactg cgctgctgag tgcttgttgc gc                       32
    573 <210> SEQ ID NO: 38
    574 <211> LENGTH: 26
E--> 575 <212> TYPE: ADN
    576 <213> ORGANISM: synthetic construct
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/673,274

Input Set : A:\98015seq.txt
Output Set: N:\CRF3\08012001\I673274.raw

578 <400> SEQUENCE: 38
579 gcaacaagca ctcagcagcg cagtca

26

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/673,274

DATE: 08/01/2001
TIME: 19:08:04

Input Set : **A:\98015seq.txt**
Output Set: **N:\CRF3\08012001\I673274.raw**

```
L:4 M:280 W: Numeric Identifier already exists, <110> found multiple times
L:5 M:280 W: Numeric Identifier already exists, <110> found multiple times
L:6 M:280 W: Numeric Identifier already exists, <110> found multiple times
L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:27 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:53 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:78 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:111 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:134 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:145 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:156 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:167 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:178 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:189 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:200 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:211 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:222 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:231 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:240 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:259 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:268 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:277 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:308 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:355 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:422 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:431 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:440 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:449 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:458 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:467 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:476 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:485 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:494 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:503 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:512 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:521 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:530 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:539 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:548 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:557 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:566 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
L:575 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:
```